RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: ___10/621,911A___
Source: ___1 FWO___
Date Processed by STIC: ___11/19/04___

ENTERED

```
          RAW SEQUENCE LISTING                    DATE: 11/19/2004
          PATENT APPLICATION:  US/10/621,911A     TIME: 10:45:21

          Input Set : A:\PTO.FG.txt
          Output Set: N:\CRF4\11192004\J621911A.raw

  3 <110> APPLICANT: SAITOU, Mitinori
  4        SURANI, Azim
  6 <120> TITLE OF INVENTION: Genes
  8 <130> FILE REFERENCE: 674558-2002
 10 <140> CURRENT APPLICATION NUMBER: 10/621,911A
 11 <141> CURRENT FILING DATE: 2003-07-17
 13 <150> PRIOR APPLICATION NUMBER: PCT/GB02/00215
 14 <151> PRIOR FILING DATE: 2002-01-18
 16 <150> PRIOR APPLICATION NUMBER: GB 0101300.2
 17 <151> PRIOR FILING DATE: 2001-01-18
 19 <160> NUMBER OF SEQ ID NOS: 26
 21 <170> SOFTWARE: SeqWin99, version 1.02
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 617
 25 <212> TYPE: DNA
 26 <213> ORGANISM: Mus musculus
 28 <400> SEQUENCE: 1
 29 gccgcagaaa gggcagaccc gcagcgcgct ccatcctttg ccctccagtg ctgcctttgc 60
 30 tccgcaccat gaaccacact tctcaagcct tcatcaccgc tgccagtgga ggacagcccc 120
 31 caaactacga aagaatcaag gaagaatatg aggtggctga gatgggggca ccgcacggat 180
 32 cggcttctgt cagaactact gtgatcaaca tgcccagaga ggtgtcggtg cctgaccatg 240
 33 tggtctggtc cctgttcaat acactcttca tgaacttctg ctgcctgggc ttcatagcct 300
 34 atgcctactc cgtgaagtct agggatcgga agatggtggg tgatgtgact ggagcccagg 360
 35 cctacgcctc cactgctaag tgcctgaaca tcagcacctt ggtcctcagc atcctgatgg 420
 36 ttgttatcac cattgttagt gtcatcatca ttgttcttaa cgctcaaaac cttcacactt 480
 37 aatagaggat tccgacttcc ggtcctgaag tgcttcaccc tccgcagctg cgtccctcct 540
 38 tgcccctccc tacacgcagg tgtaacactc atttatctat ccacagtgga ttcaataaag 600
 39 tgcacttgat aaccacc 617
 41 <210> SEQ ID NO: 2
 42 <211> LENGTH: 137
 43 <212> TYPE: PRT
 44 <213> ORGANISM: Mus musculus
 46 <400> SEQUENCE: 2
 47 Met Asn His Thr Ser Gln Ala Phe Ile Thr Ala Ala Ser Gly Gly Gln
 48 1               5                   10                  15
 50 Pro Pro Asn Tyr Glu Arg Ile Lys Glu Glu Tyr Glu Val Ala Glu Met
 51             20                  25                  30
 53 Gly Ala Pro His Gly Ser Ala Ser Val Arg Thr Thr Val Ile Asn Met
 54             35                  40                  45
 56 Pro Arg Glu Val Ser Val Pro Asp His Val Val Trp Ser Leu Phe Asn
 57         50                  55                  60
 59 Thr Leu Phe Met Asn Phe Cys Cys Leu Gly Phe Ile Ala Tyr Ala Tyr
 60 65                  70                  75                  80
```

```
62 Ser Val Lys Ser Arg Asp Arg Lys Met Val Gly Asp Val Thr Gly Ala
63                85                 90                 95
65 Gln Ala Tyr Ala Ser Thr Ala Lys Cys Leu Asn Ile Ser Thr Leu Val
66                100                105                110
68 Leu Ser Ile Leu Met Val Val Ile Thr Ile Val Ser Val Ile Ile Ile
69          115                 120                125
71 Val Leu Asn Ala Gln Asn Leu His Thr
72     130                 135
74 <210> SEQ ID NO: 3
75 <211> LENGTH: 823
76 <212> TYPE: DNA
77 <213> ORGANISM: Mus musculus
79 <400> SEQUENCE: 3
80 ggatcacaga ctgactgcta attgggtctt ggttttaggt cttttcaaag actaagcaat 60
81 cttgttccga gctagctttt gaggcttctg cccatcgcat cgccatggag gaaccatcag 120
82 agaaagtcga cccaatgaag gaccctgaaa ctcctcagaa gaaagatgaa gaggacgctt 180
83 tggatgatac agacgtccta caaccagaaa cactagtaaa ggtcatgaaa aagctaaccc 240
84 taaaccccgg tgtcaagcgg tccgcacgcc ggcgcagtct acggaaccgc attgcagccg 300
85 tacctgtgga gaacaagagt gaaaaaatcc ggagggaagt tcaaagcgcc tttcccaaga 360
86 gaagggtccg cactttgttg tcggtgctga aagaccctat agcaaagatg agaagacttg 420
87 ttcggattga gcagagacaa aaaaggctcg aaggaaatga gtttgaacgg gacagtgagc 480
88 cattcagatg tctctgcact ttctgccatt atcaaagatg ggatccctct gagaatgcga 540
89 aaatcgggaa gaattaggag cttacattgt acgctgccct ggctgtcgac gatgccgcac 600
90 agcagatgtg aaagctattt tttgtttaag attaaacttt ttctggtgct gggaaatctt 660
91 aacttgttaa cctttaaatt gtagatagga tgcacaacga tccagattta tgtgaagttt 720
92 agaagcctca agctgtgagg cccagggctg aggaataaag taaatagaat ttggagtatg 780
93 tacgttctaa tttccagaaa tttgtaataa aagcattttt gtt 823
95 <210> SEQ ID NO: 4
96 <211> LENGTH: 150
97 <212> TYPE: PRT
98 <213> ORGANISM: Mus musculus
100 <400> SEQUENCE: 4
101 Met Glu Glu Pro Ser Glu Lys Val Asp Pro Met Lys Asp Pro Glu Thr
102 1                5                 10                 15
104 Pro Gln Lys Lys Asp Glu Glu Asp Ala Leu Asp Asp Thr Asp Val Leu
105                20                 25                 30
107 Gln Pro Glu Thr Leu Val Lys Val Met Lys Lys Leu Thr Leu Asn Pro
108          35                 40                 45
110 Gly Val Lys Arg Ser Ala Arg Arg Arg Ser Leu Arg Asn Arg Ile Ala
111     50                 55                 60
113 Ala Val Pro Val Glu Asn Lys Ser Glu Lys Ile Arg Arg Glu Val Gln
114 65                 70                 75                 80
116 Ser Ala Phe Pro Lys Arg Arg Val Arg Thr Leu Leu Ser Val Leu Lys
117                85                 90                 95
119 Asp Pro Ile Ala Lys Met Arg Arg Leu Val Arg Ile Glu Gln Arg Gln
120                100                105                110
122 Lys Arg Leu Glu Gly Asn Glu Phe Glu Arg Asp Ser Glu Pro Phe Arg
123          115                120                125
125 Cys Leu Cys Thr Phe Cys His Tyr Gln Arg Trp Asp Pro Ser Glu Asn
```

Input Set : A:\PTO.FG.txt
Output Set: N:\CRF4\11192004\J621911A.raw

```
126      130                  135                 140
128 Ala Lys Ile Gly Lys Asn
129 145                 150
131 <210> SEQ ID NO: 5
132 <211> LENGTH: 4925
133 <212> TYPE: DNA
134 <213> ORGANISM: Rattus sp
136 <400> SEQUENCE: 5
137 cccccccccc cccccccccc cctccccccc ccccccacctc cgacgtatga tggctcctag 60
138 acgcaacacg aagcggactc cccgcatcat tcacgtagac ccgccttctg ctttccctgt 120
139 cggggttttg ggaagcccgg cggccctctc ttctcacctt gctccactag cacgcggctg 180
140 ttttcactga gcccagcact ggctaagtgg agcaccagga gtttcaggct atccttcaga 240
141 gggcaaggtg tagtccatgg tgggctacag gagaccctct ctctccgtga gtacagagag 300
142 gcaaacccaa gccagacagg ggtgatgatt aggaacatac cttcgtcggg gagaaaatac 360
143 cggttcatat aggaataaga ggaaccagga ggtagttaag gctgtggtgt ctggttgcgg 420
144 ggttttgac tctcaacaac cacgttcaga acgtgctgag tttttatgat ggtgtagaat 480
145 ttccttatca gcaattggtc tccgcggtgt ttctttttct tttttaattt tttaagtata 540
146 atttggtgtt tgaagcaact gtacttggac tagaactccc tgtgtaatcc agaatggaat 600
147 cccaaatcct aggattaaag gtttttagtgg gctgcagtgt ggggtggggg ttgttttgat 660
148 tacgttgtag cccaggctgg gctcaatctc aatcctcctg cctctgcctt ctaaacgcta 720
149 ggattaaaag tgctgcgcca tgatcctgct gtagctttat ttttatttat ttatttattt 780
150 attttggctc tttttttttg gagctgggga ccgaaccgag ggccttgtgc ttcctaggca 840
151 agcgctctac cactgagcta aatccccaac cccagtgtag ctttattttt aagaacagga 900
152 gtcttgtttc tcaaaacagt ttctctgtag ccctggttgt cctggaactc cgtaaaccag 960
153 gctggtttgg gactctgcct ttaaaacact gggactaaag gcggtaccac ctccgtgggc 1020
154 tacaccggaa tcttttaagc ttcatttgaa ccggggcttt ttctttttct cacccacttt 1080
155 ctggaagcga ttttcctgct aaatttccat tcctggtaaa tgactctgag gggaaatagg 1140
156 aacccagaat agattgagcc ggggggctacc tgggaccccg cactccccac cccccagccg 1200
157 ctgttgaagc tctttgcctg aggggcctcc gggtttgata cctcctagca ctccgggctg 1260
158 agggcgtggc tcgggaggag ccattccttt ggagaggaaa acaactgctg gccttgaatc 1320
159 tgccctaata cctgacagtt acatgggacc tccttatttc cacaggattc tttagtcttt 1380
160 gtttgggaga ttttcaaatc ttgagactgc tcaacccttc ctggcctaac actcacaagg 1440
161 ccaggctaga cccaaattct gtcaacccct tctgtgtcca aaacggtggg tggctagctg 1500
162 gctcacccctt ggtgtcactt tgctttaaca ttcggaaaag ttgtggtaag tttcctgtat 1560
163 aaaataggac catctactgg gtgtggtccc atgtaaagca aggttggttt cccaaaatac 1620
164 cctgtttaca tagatgtccg gaagcattgg agcaggtcaa ttagatttag gtggaaacag 1680
165 cctgtttttg gaaagctttc cagggcggaa aatgaaccca gaggcactat tgggcaagcc 1740
166 ctccggctaa gcaacacaat tggctgcagg ggtctctgga agaggtgtga gacaagagag 1800
167 aatatgcagg tttcaggacc tctgaactag agttaggctg ctgtaacatt gtaacattgc 1860
168 tgtaagcaga acagcccatg gtaagaagct cagtggatct ctacaaacac taggatatct 1920
169 gctcagggtt tatgaccagg ccctgtgcat atggtttgct tcttgttggc ccctctcttg 1980
170 aagaggggtg attatctgtt acccacttcc ttgtttctct ggggtattac cttgcaaaat 2040
171 gcaaaatgat atacttcact aatgtctcca tcttctgttt cagaaatcct acaaccagaa 2100
172 acactagtaa aggtcatgaa aaagctaacc ctgaacccca gtgccaagcc gacaaaatat 2160
173 catcgtcgtc aaagggttcg tctccaggtt aagagccagc ctgtggagaa cagaagtgaa 2220
174 agaatcatga gggaagttca aagcgccttt cccaggagaa gggtccgcac tctgttgtcc 2280
175 gtgctgaaag accccatagc aaggatgaga agatttgttc gggtgagttg cgtttgtggg 2340
176 cggggcatag atctaagagc aactctagcc tcaggaatgg cacctaggtt aaacagggaa 2400
177 tgtagacaag gatagtgact acctgtgatt cccagctcaa gaaaacaagc tccaaggcta 2460
```

```
178 tcctctactg cgcagtctga agctggccag agctatatgc aaattgataa gtcagtataa 2520
179 catttatttt tggattttca gactccctcc ccatagtcca aactggccct ccagttcagt 2580
180 ccacggtcct gcttcttccc cggtgctagg cttttgagtg ataaggctga cttagactgg 2640
181 atctcagagc tgaagtggac ctgttagtct ttgtagacca ggctggggtg gtttctgctt 2700
182 tctcagcgcc tagctcacat agtaggcatt ttaactttgt cttaatagta atttgagtaa 2760
183 ttttgttttt ctcttgaaga ttgagcagag acaaagacag cttgaaggaa atgaggtaaa 2820
184 tgcatatgga tgggtagggt gtctatggat gggtagggtg tcttgttttt actgtttcct 2880
185 tagacaagga gtgtgtatgt ggagagttac cttctcaaca cagggaatct ggttattaaa 2940
186 gcagtacttt aaaaataaat aaaataaata aaataaaaat aaagcagtag aaggggattt 3000
187 acatttcttt tgagttgcaa tatcctgatt aacatttttc tttcagagac gagatgagcc 3060
188 attcagatgt ctctgcactt tctgccatta tcagagatgg gatccttctg agaatgctaa 3120
189 aatcgggcag aaccagaaga attagggcag tttgaattgt acaccgtcct tgccgttaac 3180
190 ggtgccatgc agcagatgtg aaagctgttt ttttgtttaa gattaaactt ttcttggtgc 3240
191 tggggaaatc tcttctaatt gctaaccttt aaattatata ggatgtgtga catttggatt 3300
192 catgggaatg acagatttac ccaagaattg agcatgagtc aaagcctggt agtttgattt 3360
193 agaaggtaat tggaataaat ctttttattt tagattttct agtttgcaga gaaatttgta 3420
194 ataaaggcaa atttgttatc tttaataaat acagaacaga ttagaatgag ccattggaga 3480
195 tgggggactc gttttttaca ggtgcatgtg tgggtgtgtg atgttcagag ttcaatgtgt 3540
196 gctaccctgt atttctgctt gaggcaaggt ctccatgagg cctagctggt ctaactcctg 3600
197 gtcctgcctt ttgtttttccc ctgagttttg acaccatagg cttgtcggca agatctggaa 3660
198 gaggcttgat gtttgtgttt gtgctgtgta ataaacaatt ggttgacata ttcctaaagt 3720
199 gtggcactgt attgacctgt ctgtctcatg aggaagttaa tgaccggagc ataattgtat 3780
200 gctttatttc ctgagagaag tgtcaggaaa ggaggagtta ggaagaaagc cccaggctgg 3840
201 ggttaagagc actggctgct tttccagagg tcctgagttc aattcccagc aatcacctgg 3900
202 tggctcccga acatctgtaa caggatccaa tgccctcttt tggtgtgtct aagaactccc 3960
203 taggcatgca gaggattttt gtttttgttt ttttttttttt ttttttttttt ttcgtttttt 4020
204 tcagagctgg ggaaccgaac ccagggcctt gcgcttgcta agcaagcgct ctaccactga 4080
205 gctaaatccc caaccccctac aatggccttt ttctacctgc ttttgaatta tcaataaaag 4140
206 actggggcaa aagaaaggct ggagtgaatg agagagaaca tgtgaagagt aaatgagaga 4200
207 gagcatgagg gaatgaatga gagagtgaat gtgagaacga atgtgagagc gagtgagaga 4260
208 acatgagaag aacacgttaa gagtgagtga agagagaatg tgaggtgtgt atgaagattg 4320
209 tgtgtggggt tggggattta gctcagtggt agagtgcttg cctaggaagc acaaggccct 4380
210 gggttcggtc cccagctcca aaaaaaagac ccaaaaaaaa aaaaaaaaaa aaagattgtg 4440
211 tgtgtgtgtg aaaggagagt gcatgtggtg tgtgtgagat atgtgcaagg tgtgtatcaa 4500
212 gagtgtgtgt gagagtgaaa gggtaatgaa cagaggtgtg catgagcgtg ggagtttgag 4560
213 aaaagaaaac agcaataaaa aaaaaagcag agtgcacgag agaatgcaga gtgtgtgcaa 4620
214 cctcaagctg agacagagac agagagaaag agagagagag agagagactt taagccttga 4680
215 aattacctgt cagtttgtac ccaaatagta gtctgtgtat atttattttg agccttccag 4740
216 atccctgctt ccagtggaga actctgattc tatgttgagg ctggaccctg gcaatagtgg 4800
217 gcttcttgaa aaatagtcaa aggaaacagt gctacaccat ggacttaagc ctttagactc 4860
218 agttctggct tcaagagcag ctgtcagaaa ataagtgatg aactacttgc agtcgaactc 4920
219 gaatc 4925
221 <210> SEQ ID NO: 6
222 <211> LENGTH: 1444
223 <212> TYPE: DNA
224 <213> ORGANISM: Rattus sp
226 <400> SEQUENCE: 6
227 ccaggattca gacgagctag gcctcatgca tggagacctt gcctcaagca gaaataaaca 60
228 gggtagcaca cattgaactc tgaacatcac gagtgtgcac acacccacac atgcatctgt 120
```

```
229 aaaaaacgag tccccatctc caatggctcg ttctaatctg ttctgtgtat ttattaaaga 180
230 taacaaattt gcctctatta caaatttctc tgcaaactag aaaatctaaa ataaaagatc 240
231 tattccaatt accttctaaa tcaaactacc gggctttgac tcatgctcaa ttcttgggta 300
232 aatctgtcat tcccatgaat ccaaatgtca cacatcctat ataatttaaa ggttagcaag 360
233 tagagatttc cccagcacca agaaaagttt aatcttaaac aaaaaaacag ctttcacatc 420
234 tgctgcatgg caccgttaac ggcaaggaca gtgtatgatt caaactgccc taattcttct 480
235 ggttctgccc aattttagca ttctcagaag gatcccatct ctgataatgg cagaaagtac 540
236 agagacatct gaatggctca actcttctct catttccttc aagctgtctt tgtctctgct 600
237 caatccgaac aaatcttctc atccttgcta tggggtcttt cagcaccgac aacagtgtgc 660
238 ggacccttct cttgggaaag gcgctttgaa ctcccctcat gattctttca cttctgttct 720
239 ccacaggctg gctcttaatc tggagacgaa ccctttgacg aagatgatat tttggccgat 780
240 tgagatagaa tatcaaaaca acatttaaca tttaaataac ttaacgatat acacacccttt 840
241 ttttttttcca cctccccaca cagacaaaaa acaaccctat ttttttctttta caaccccgcc 900
242 taagcaagcg aagcattagt aactgaccaa tcatagaaag gaaacaccac cagaccacat 960
243 caaataaaat aaaatcaccg cccaacccca cccctataaa aaacccgccg accacaccac 1020
244 atatactccc ccccccccgc accatcacta catcaccctc tccacccatt cccacctccc 1080
245 cccccaacat taaccccacc ccatcacgga aaccccaac accaacaaat aaattagaca 1140
246 catcgcatta cataaattga cacaagaccc accccaaaag agcagcaaag attagagcca 1200
247 catcctcggc ccaacacaat acactcaacc tgcatagtat ctatctccac cccaacctag 1260
248 aaacaaaaat ctaatcagca ccaggcaccc aagtatcacg cacactcaaa aacatacccca 1320
249 ccaattaaac acgccccacc cacccaacaa cccacccgcc tgacaacaca cttcggaact 1380
250 accctcaaca tcaccaaaag caatcgcaag ttacgatgac tccaaccacc tcactctctc 1440
251 attg 1444
253 <210> SEQ ID NO: 7
254 <211> LENGTH: 7656
255 <212> TYPE: DNA
256 <213> ORGANISM: Rattus sp
258 <220> FEATURE:
259 <221> NAME/KEY: misc_feature
260 <222> LOCATION: (7471)..(7471)
261 <223> OTHER INFORMATION: "n" is an unknown nucleotide
263 <220> FEATURE:
264 <221> NAME/KEY: misc_feature
265 <222> LOCATION: (7554)..(7554)
266 <223> OTHER INFORMATION: "n" is an unknown nucleotide
268 <220> FEATURE:
269 <221> NAME/KEY: misc_feature
270 <222> LOCATION: (7608)..(7608)
271 <223> OTHER INFORMATION: "n" is an unknown nucleotide
273 <400> SEQUENCE: 7
274 ctgcaagtag ttcatcattt acagatcaaa agaaagaaga ataaaaaaac aaggtgtcat 60
275 gatccctcca aaagagtgga acacttcaac tgccagatcc aagatactga aatgggtagc 120
276 atgctggaga aagaattcaa aagttaggta gagaatctgg ttgagcagag cacttgcttt 180
277 tcttccagag gatctgagtt caagtcccag gacctatatc acagttttct gtaactctag 240
278 ctccagaggg tctgacactt ctgttcactg tgggcacctg cattcacaga caaacataaa 300
279 gtagttcatc acccttttca cagaaaaccc acagcatgtg aggaaatccg ggtctctgcg 360
280 caatgccccc acagcagaag ggggggagctg gagagatggt tcatctgtta gcccatttat 420
281 tgctcttgaa gagaacccag ggtcatccat agcacccata gcagctcaca accatctcca 480
282 gttccaggag atccaatgcc ctgttgtgac ctcaggtacc aggcatacac aatgaacctg 540
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 11/19/2004
PATENT APPLICATION: US/10/621,911A   TIME: 10:45:22

Input Set : A:\PTO.FG.txt
Output Set: N:\CRF4\11192004\J621911A.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 7471,7554,7608
Seq#:8; N Pos. 2115,2142,2143,2146

VERIFICATION SUMMARY   DATE: 11/19/2004
PATENT APPLICATION: **US/10/621,911A**   TIME: 10:45:22

Input Set : **A:\PTO.FG.txt**
Output Set: **N:\CRF4\11192004\J621911A.raw**

L:398 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:7440
M:341 Repeated in SeqNo=7
L:464 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:2100